What is this all about

- Human pose detection and tracking
- Human recognition
- Applications in:
  - Human robot interaction
  - Medical field
  - Entertainment
  - Automotive
  - ...
- Create a "People Library"
- Triggered by work of J. Shotton MSR
- OpenNI NITE library
- Cheap (RGB-D) HW:
  - Kinect
  - Xtion Pro
  - Dinast
  - Raytrix
  - SoftKinetic
  - BrainVision
  - ...

People Library / Talk ICRA2013 / Point Cloud Library
Motivation

- Moving platform
  - No BG subtraction
  - Separate segmentation and pose estimation?
- Cluttered scenes
- SDK gives only skeleton
- Processing power
- Open Source, modular
First implementation of people library
Plug in of extra "detectors"
Open Source
RDF based
Adds bio-mechanical knowledge
Workflow

- **Off-line, Training:**
  - Build 'human' database
  - Synthesize training data
  - Train RDF trees on this data

- **On-line:**
  - Run labeling based on RDF
  - Search within kinematic constraints
  - Evaluation global consistency
  - Temporal tracking
Off-line, Training:
- Build ‘human’ database
- Synthesize training data
- Train trees on this data

On-line:
- Run labeling based on RDF
- Search within kinematic constraints
- Evaluation global consistency
- Temporal tracking
Off-line, Training:
- Build 'human' database
- Synthesize training data
- Train trees on this data

On-line:
- Run labeling based on RDF
- Search within kinematic constraints
- Evaluation global consistency
- Temporal tracking
Off-line, Training:
- Build 'human' database
- Synthesize training data
- Train trees on this data

On-line:
- Run labeling based on RDF
- Search within kinematic constraints
- Evaluation global consistency
- Temporal tracking
Workflow

- **Off-line, Training:**
  - Build 'human' database
  - Synthesize training data
  - Train trees on this data

- **On-line:**
  - Run labeling based on RDF
  - Search within kinematic constraints
  - Evaluation global consistency
  - Temporal tracking
Workflow

- **Off-line, Training:**
  - Build ‘human’ database
  - Synthesize training data
  - Train trees on this data

- **On-line:**
  - Run labeling based on RDF
  - Search within kinematic constraints
  - Evaluation global consistency
  - Temporal tracking
MapReduce problem
Solved in Hadoop
3 Trees, 20 levels
2000 features/pixel
2000 pixels/image
290GB features for single person in 80k poses
7 day training
Training Problems

Training

- 1 Person is limited
- 80k poses is not enough
- > more data
- **Label PointCloud - RGB-D**
- Search feasible blobs
- Build local consistency map
- Find local consistency tree
- 3D Grabcut
- Iterate again
Runtime Workflow

- Label PointCloud - RGB-D
- Search feasible blobs
- Build local consistency map
- Find local consistency tree
- 3D Grabcut
- Iterate again
Label PointCloud - RGB-D
Search feasible blobs
Build local consistency map
Find local consistency tree
3D Grabcut
Iterate again
Label PointCloud - RGB-D
Search feasible blobs
Build local consistency map
Find local consistency tree
3D Grabcut
Iterate again
Workflow

- Label PointCloud - RGB-D
- Search feasible blobs
- Build local consistency map
- Find local consistency tree
- 3D Grabcut
- Iterate again
Workflow

First Iteration
- Noisy skeletons
- Clusters
- Pixel labelling

Second Iteration
- Final skeleton
- Clusters
- 3D Segmentation
- Refined label

$t-1$, $t$, $t+1$
Examples

Negative
Clutter
Clutter
On PC:

- Fermi, Kepler GPU
- CUDA Toolkit 4.x
- RGBD camera, organised and projectable
- 20-100fps

On Android:

- Tegra 4 chipset
- 10-20 fps
- Only first iteration
API

Input:
- (RGB)-D

Output (per person):
- BG/FG Segmentation
- Skeleton
- Pixel labeling
Overview: Prob

Body Pose + Parameter Estimation
\[ p(X_C \mid \{c, \ l, \ x\}) \rightarrow p(X_R \mid \{X_C\}) \]

Body Part Evaluation and Pose Estimation
\[ p(c \mid \{l, \ x\}) \rightarrow p(X_C \mid \{c, \ l, \ x\}) \]

Pixel-Body Part Labeling
\[ \{l, \ x\} \rightarrow p(c \mid \{l, \ x\}) \]

RGBD Sensor Organized PointCloud XYZRGB

Time t

Time t+1
Overview: Prob: Result
Data Set: Vicon

Evaluation

- 60 people
- 30 min motions
  - Static poses
  - Free motions
  - Interaction with objects (cooking tasks)
- 8 Primesense camera’s, 5 GigE camera’s
- Vicon mocap, 61 markers on person
- > Will become public available
Data Set: Real life

Evaluation

- 40 people
- PR2 in office environment
- Sitting, Walking, Interaction
Surgeon tracking

What next?

- Laparoscopic surgeries
- Tracking pose
- Ergonomics, control, ...
Going embedded

What next?

- Android version
- Available in pcl/trunk
- Early result, new adjusted training needed
People

Koen Buys (KU Leuven)
Anatoly Baksheev (Itseez)
Cedric Cagniart (TUM)
Caroline Pantofaru (Willow Garage)
Radu B. Rusu (Open Perception)
...

Who
Acknowledgements

Sponsored by:

▶ University of Leuven
▶ TU Munich
▶ Willow Garage
▶ Itseez
▶ Nvidia
▶ Amazon Web Services
Questions?

buys dot koen at gmail dot com